



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGATTATCCTACTTTACTTTGGCTTCTTTCATGT
1 M D Y P T L L L A L L H V
601 ATACAGAGCTCTATGTGAAGAGGTGCTTTGGCATAACATCAGTTCCTTTGCCGAGAACATGTCTCTAGAATGTGTGTATCCATCAATGGGCATCTTAACA
13 Y R A L C E E V L W H T S V P F A E N M S L E C V Y P S M G I L T
701 CAGGTGGAGTGGTTCAAGATCGGGACCCAGCAGGATCCATAGCCATTTTCAGCCCTACTCATGGCATGGTCATAAGGAAGCCCTATGCTGAGAGGGTTT
47 Q V E W F K I G T Q Q D S I A I F S P T H G M V I R K P Y A E R V
801 ACTTTTTGAATTCAACGATGGCTTCCAATAACATGACTTTTTCTTCGGAATGCCTCTGAAGATGATTTGGCTACTATTCTGCTCTTTACACTTA
80 Y F L N S T M A S N N M T L F F R N A S E D D V G Y Y S C S L Y T Y
901 CCCACAGGAACTTGGCAGAAGGTGATACAGGTGGTTCAGTCAGATAGTTTTGAGGCAGCTGTGCCATCAAATAGCCACATTTTTCGGAACCTGGAAG
113 P Q G T W Q K V I Q V V Q S D S F E A A V P S N S H I V S E P G K
1001 AATGTCACACTCACTTGTGAGCCTCAGATGACGTGGCCTGTGCAGGCAGTGAAGTGGGAAAAGATCCAGCCCCGTGATCGACCTCTTAACCTACTGCA
147 N V T L T C Q P Q M T W P V Q A V R W E K I Q P R Q I D L L T Y C
1101 ACTTGGTCCATGGCAGAAATTTACCTCCAAGTCCCAAGACAATAGTGAGCAACTGCAGCCACGGAAGGTGGAGGCTCATCGTCATCCCCGATGTCAC
180 N L V H G R N F T S K F P R Q I V S N C S H G R W S V I V I P D V T
1201 AGTCTCAGACTCGGGCTTACCGCTGCTACTTGCAGGCCAGCGCAGGAGAAAACGAAACCTTCGTGATGAGATTGACTGTAGCCGAGGGTAAAACCGAT
213 V S D S G L Y R C Y L Q A S A G E N E T F V M R L T V A E G K T D
1301 AACCAATATACCCTTTGTGGCTGGAGGGACAGTTTTATTGTTGTTGTTGTTATCTCAATTACCACCATCATTGTCATTTTCTTAAACAGAAGGAGAA
247 N Q Y T L F V A G G T V L L L L F V I S I T T I I V I F L N R R R
1401 GGAGAGAGAGAAGAGATCTATTTACAGAGTCTGGGATACACAGAAGGCACCAATAACTATAGAAGTCCCATCTCTACCAGTCAACCTACCAATCAATC
280 R R E R R D L F T E S W D T Q K A P N N Y R S P I S T S Q P T N Q S

NheI (1598)

1501 CATGGATGATACAAGAGAGGATATTTATGTCAACTATCCAACCTTCTCTCGCAGACCAAGACTAGAGTTAAGCTTATTCTTGACATGAGTGCATTAGC
313 M D D T R E D I Y V N Y P T F S R R P K T R V •
1601 TAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATT
1701 GCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGGAGGTGGGAGGTTTTTT
1801 AAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGA
1901 GGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCC
2001 CAAGTGTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTTCCCTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCAT
2101 TGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAAC
2201 CTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGC
141 • N R T Y K L P I L E E I T T K V L K G
2301 CATTCACTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTC
121 N M E I L V F C D P A Y D S I L E R C M G C P S V V R I S R D V E
2401 ATCAGAGTAGGGGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCCAGCACAGACAGTGACC
88 D S Y P H R V A V I T D F D K Q G N S V A S G I A I A E A C V T V
2501 CTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCTCCTCATAGAGCATGG
54 R G I Y A E I H V A S I I E G T K T R I A A G V H H K N D E Y L M T
2601 TGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGC
21 I K E T A V E V L E L D Q Q S I N F T K M
2701 CGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACTCCCA
2801 CCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACG
2901 TCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTA
3001 ATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGT
3101 ACTTGGCATATGATACACTTGTGACTGCAAGTGGGCGAGTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATG

3201 GGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAITAA
3301 GAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA
3401 AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCC
3501 TGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCG
3601 CTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGTGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
3701 TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTA
3801 GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACACCGCTGGTAGCGG
3901 TGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAA
4001 AACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTT
4101 TTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGA
4201 ACATTTCTCTATCGAA